At page 16, line 15 to page 17, line 4:

A full length transcript is obtained by using calf intestinal alkaline phosphatase followed by tobacco nucleotide acid pyro-phosphatase (TAP) with an RNA ligase reaction that links decapped mRNA to a specific oligoribonucleotide (38-mer). This 38-mer would contain a T7-promoter. Reverse transcriptase is then used to extend full length cDNA using one or more of the following three primers with 3' ends complementary to the three possible stop codons TAA, TGA, TAG):

5' GTGCAGNNNNNNNNNNNNTTA [SEQ ID NO:22]

5' GTGCAGNNNNNNNNNNNNCTA [SEQ ID NO:23]

and/or

5' GTGCAGNNNNNNNNNNNNTCA [SEQ ID NO:24]

RnaseH then prime the second strand cDNA with the 38 mer complement 3' overhang at one end of a double stranded fragment which contains the NEB IMPACT (Intein Mediated Purification with an Affinity Chitin-binding Tag). At the other end of the fragment is another BsgI site next to first intein codon (TGC = cys). The fragment can then be cleaved with BsgI, which creates two base 3' overhangs at the stop codon (TAR or TGA) and cys codon:

5' ATGNNN...NNNTA CTGCTTTGCCAAGGGTACCAATG [SEQ ID NO:25]

TACNNN...NNN ATGACGAAACGGTTCCCATGGTTAC 5' [SEQ ID NO:26]

and/or

5' ATGNNN...NNNTG CTGCTTTGCCAAGGGTACCAATG [SEQ ID NO:27]

TACNNN...NNN CTGACGAAACGGTTCCCATGGTTAC 5' [SEQ ID NO:28]

At page 63, lines 17 to 23:

In order to evaluate the method, molecules of a known DNA sequence were first cast into a polyacrylamide gel matrix. The oligonucleotide sequencing primer RMGP1-R (5' - gcc cgg tct cga gcg tct gtt ta [SEQ ID NO:29]) was annealed to the oligonucleotide puc514c (Q - 5' tcggcc aacgcgcggg gagaggcggt ttgcgtatca g **taaacagac gctcgagacc gggc** [SEQ ID NO:30] (sample 1)) or to the oligonucleotide puc234t (Q - 5' cccagt cacgacgttg taaaacgacg gccagtgtcg a **taaacagac gctcgagacc gggc** [SEQ ID NO:31] (sample 2). The bolded sequences denote the sequences to which the sequencing primer anneals, and Q indicates an ACRYDITE modification.

At page 68, line 16 to page 69, line 2:

Another modification would address the difficulty, encountered in many methods, of sequencing past long repeating stretches. If it is known that a given array contains many such sequences, one may include a defined regimen (for example, halfway through the whole sequence) of deoxy- and dideoxynucleotides to reduce out-of-phase templates. That is, if one knows he or she is sequencing through a repeat of, for example, AC dinucleotides, one may reduce the number of out-of-phase molecules by following a dATP addition with a ddATP addition. Only those molecules which failed to incorporate the deoxy- form of the nucleotide will be available to incorporate the dideoxy- form, leading to chain termination and reduction of that source of background. Clearly, similar regimens may be devised for repeats involving more than two nucleotides. It should be noted that the strategy is not limited to repeats and may be used to extend read length in any situation where most of the sequences in the array have a block of sequence part of the way through the target sequence which is known. For example, in an array of targets, most having the unique sequence ACGTA [SEQ ID NO:32] at the same distance

from the primer, one may reduce the number of out-of-phase molecules by following a dATP addition with a ddATP, ddGTP, and ddTTP addition, then dCTP followed by ddATP, ddCTP, and ddTTP addition.

At page 71, lines 13 to 19:

In order to evaluate the method, molecules of a known DNA sequence were first cast into a polyacrylamide gel matrix. The oligonucleotide sequencing primer RMGP1-R (5' - gcc cgg tct cga gcg tct gtt ta [SEQ ID NO:29]) was annealed to the oligonucleotide puc514c (Q - 5' tcggcc aacgcgcggg gagaggcggt ttgcgtatca g **taaacagac gctcgagacc gggc** [SEQ ID NO:30] (sample 1)) or to the oligonucleotide puc234t (Q - 5' cccagt cacgacgttg taaaacgacg gccagtgtcg a **taaacagac gctcgagacc gggc** [SEQ ID NO:31] (sample 2). The bolded sequences denote the sequences to which the sequencing primer anneals, and Q indicates an ACRYDITE modification.

At page 76, line 28 to page 77, line 8:

MmeI is a restriction endonuclease having the property of cleaving at a site remote from its recognition site, TCCGAC [SEQ ID NO:33]. Heterogeneous pools of primers are constructed that comprise (from 5' to 3') a sequence shared by all members of the pool, the MmeI recognition site, and a variable region. The variable region may comprise either a fully-randomized sequence (e.g. all possible hexamers) or a selected pool of sequences (e.g. variations on a particular protein-binding, or other, functional sequence motif). If the variable sequence is random, the length of the randomized sequence determines the sequence complexity of the pool. For example, randomization of a hexameric sequence at the 3' ends of the primers results in a pool comprising 4,096 distinct sequence combinations. Examples of two such mixed populations of oligonucleotides (in this case, 32-mers) are primer pools 1s and 2s, below: